NOVEL PLANT PROMOTERS FOR USE IN EARLY SEED DEVELOPMENT

This application claims the benefit of U.S. Provisional Application 60/458,828 filed March 28, 2003, which is incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to the field of plant genetics. More specifically, the present invention relates to early seed development gene expression. The present invention provides promoters capable of transcribing heterologous nucleic acid sequences in seeds, and methods of modifying, producing, and using the same. The compositions comprise novel nucleotide sequences for plant promoters, more particularly the seed coat promoter from *Arabidopsis*, pBAN; and the early seed promoters from *Arabidopsis*, p26, p63, and p63tr that are useful for expression of transgenes of agronomic importance in crop plants.

BACKGROUND OF THE INVENTION

A large number of genes are known which are expressed only in developing seeds, or are expressed in developing seeds at much higher levels than in any other organ or tissue type. Much of the information about seed-specific gene expression has been derived from studies of genes encoding storage proteins (reviewed by Bevan *et al.*, *Phil. Trans. Royal Soc. Lond. Biol. Sci.*, 342:209-215 (1993)). For instance, DNA sequences that confer embryo-specific expression by the soybean conglycinin promoter in transgenic plants have been identified (Chen *et al.*, *EMBO J.*, 6:3559-3564 (1988)). Similarly, the storage protein napin is one of the major protein components of *Brassica napus* L. (oilseed rape) seeds. The 5' regulatory region from napin has been published (Kridl *et al.*, *Seed Sci. Res.*, 1:209-219 (1991)). A 152 bp fragment from the napin promoter directed strong expression of the ß-glucuronidase reporter gene in mature tobacco seeds (Stalberg *et al.*, *Transgenic Research*, 7(3):165-172 (1998)). The napin promoter has been used to control expression of genes in transgenic plants designed to produce novel fatty acids (*e.g.*, Voelker *et al.*, *Plant Journal*, 9:229-241 (1996)). However, because storage lipid accumulation begins substantially before the maximal level of expression of the napin or other storage protein genes is reached (Post-Beittenmiller *et al.*, in *Control of Plant Gene Expression*. Verma, D. P. (ed.) Telford Press, pp. 157-174 (1992)), the promoters of storage protein genes may not always be preferred for controlling expression of genes related to oil accumulation in plant seeds.

Current technology permits the transformation of plants with heterologous genes. The expression of these genes is either ubiquitous if the promoter is constitutive, or is regulated in a temporal or spatial manner if the promoter is stage- or tissue-specific. Continuous expression precludes production at particular stages or in specific tissues, and can adversely

affect yield due to increased energy demands associated with prolonged synthesis of the product. Tissue- or stage-specific expression permits greater control over the temporal and spatial accumulation of desired products. Thus, promoter sequences that control the expression of desired genes in a tissue-specific, stage-specific manner that can be employed in recombinant constructs for the transformation of plants, and that would facilitate greater control over the location, timing, and duration of expression of introduced genes and reduce the possibility of deleterious effects on overall plant growth, are highly desirable.

For production of transgenic plants with various desired characteristics, it would be advantageous to have a variety of promoters to provide gene expression such that a gene is transcribed efficiently in the amount necessary to produce the desired effect. The commercial development of genetically improved germplasm has also advanced to the stage of introducing multiple traits into crop plants, often referred to as a gene stacking approach. In this approach, multiple genes conferring different characteristics of interest can be introduced into a plant. It is often desired when introducing multiple genes into a plant that each gene is modulated or controlled for optimal expression, leading to a requirement for diverse regulatory elements. In light of these and other considerations, it is apparent that optimal control of gene expression and regulatory element diversity are important in plant biotechnology.

SUMMARY OF THE INVENTION

In the present invention, we provide DNA molecules that encode transcription regulatory regions useful in driving expression of selected polynucleotide molecules at specific times and in specific tissues in plant cells.

The present invention provides and describes compositions and methods for regulating expression of heterologous polynucleotide molecules in a plant. The compositions comprise novel nucleotide sequences for plant promoters, more particularly the seed coat promoter from *Arabidopsis*, pBAN; and the early seed promoters from *Arabidopsis*, p26, p63 and p63tr.

In one embodiment, the present invention provides a promoter comprising a polynucleotide sequence selected from the group of polynucleotide sequences consisting of: a polynucleotide sequence from the group of polynucleotide sequences consisting essentially of SEQ ID NO: 1 and SEQ ID NO: 4, a polynucleotide sequence substantially homologous to SEQ ID NO: 2 or any fragments or regions thereof, and a polynucleotide sequence comprising SEQ ID NO: 3.

In another embodiment, the present invention provides a plant expression construct comprising a promoter comprising a polynucleotide sequence selected from the group of polynucleotide sequences consisting of: a polynucleotide sequence from the group of polynucleotide sequences consisting essentially of SEQ ID NO: 1 and SEQ ID NO: 4, a polynucleotide sequence substantially homologous to SEQ ID NO: 2 or any fragments or regions thereof, and a polynucleotide sequence comprising SEQ ID NO: 3, wherein said promoter is operably linked to a transcribable polynucleotide molecule. In a preferred embodiment, the transcribable polynucleotide molecule is a gene of agrinomic interest. In a preferred embodiment, the transcribable polynucleotide molecule is a marker gene.

In yet another embodiment, the present invention provides a transgenic seed-producing dicotyledonous plant stably transformed with a plant expression construct comprising a promoter comprising a polynucleotide sequence selected from the group of polynucleotide sequences consisting of: a polynucleotide sequence from the group consisting essentially of SEQ ID NO: 1 and SEQ ID NO: 4, a polynucleotide sequence substantially homologous to SEQ ID NO: 2 or any fragments or regions thereof, and a polynucleotide sequence comprising SEQ ID NO: 3, wherein said promoter is operably linked to a transcribable polynucleotide molecule operably linked to a 3' transcription termination polynucleotide molecule. In a preferred embodiment, the dicotyledonous plant is selected from the group consisting of tobacco, tomato, potato, peanut, soybean, cotton, canola, rapeseed, safflower, flax, sugarbeet, *Arabidopsis*, *Brassica*, sunflower and alfalfa. In a more preferred embodiment, the transgenic dicotyledonous plant has seed with altered protein content. In a more preferred embodiment, the transgenic dicotyledonous plant has seed with altered oil content. In a more preferred embodiment, the transgenic dicotyledonous plant has seed with altered micronutrient content. In a more preferred embodiment, the present invention provides seed, oil, or meal of a transgenic dicotyledonous plant.

In another embodiment, the present invention provides a method of making a vegetable oil and meal, comprising the steps of incorporating into the genome of a dicotyledonous seed-producing plant a promoter of the present invention operably linked to a transcribable polynucleotide molecule conferring altered oil content, growing the dicotyledonous plant to produce seed, and extracting the oil from the seed to produce extracted oil and meal.

The foregoing and other aspects of the present invention will become more apparent from the following detailed description and accompanying drawings.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

SEQ ID NO: 1 represents a nucleic acid molecule encoding a pBAN promoter.

SEQ ID NO: 2 represents a nucleic acid molecule encoding a p26 promoter.

SEQ ID NO: 3 represents a nucleic acid molecule encoding a p63 promoter.

SEQ ID NO: 4 represents a nucleic acid molecule encoding a p63tr promoter.

SEQ ID NO: 5 represents the P1 clone:MQL5gi|3702735|dbj|AB018117.1| from chromosome 5 of *Arabidopsis thaliana*.

SEQ ID NO: 6 represents the BAC clone T13M11.

SEQ ID NO: 7 is a primer sequence for PCR amplification identified as Clone 26 GSP1.

SEQ ID NO: 8 is a primer sequence for PCR amplification identified as Clone 26 GSP2.

SEQ ID NO: 9 is a primer sequence for PCR amplification identified as BAN+1500.

SEQ ID NO: 10 is a primer sequence for PCR amplification identified as pBAN GSP1.

SEQ ID NO: 11 is a primer sequence for PCR amplification identified as BAN-Nco.

SEQ ID NO: 12 is a primer sequence for PCR amplification identified as p63-Nco.

SEQ ID NO: 13 is a primer sequence for PCR amplification identified as p63-fwd3.

SEQ ID NO: 14 is a primer sequence for PCR amplification identified as GUS 5'.

SEQ ID NO: 15 is a primer sequence for PCR amplification identified as GUS 3'.

SEQ ID NO: 16 is a primer sequence for PCR amplification identified as CP4-Dra.

SEQ ID NO: 17 is a primer sequence for PCR amplification identified as CP4-Kpn.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a schematic representation of pMON69802.

Figure 2 is a schematic representation of pMON69804.

Figure 3 is a schematic representation of pMON69815.

Figure 4 is a schematic representation of pMON69812.

Figure 5 is a schematic representation of pMON82350.

Figure 6 is a schematic representation of pMON65422.

Figure 7 is a schematic representation of pMON65428.

DETAILED DESCRIPTION OF THE INVENTION

The following definitions and methods are provided to better define the present invention and to guide those of ordinary skill in the art in the practice of the present invention.

Unless otherwise noted, terms are to be understood according to conventional usage by those of ordinary skill in the relevant art.

As used herein, the phrase "polynucleotide molecule" refers to the single- or double-stranded DNA or RNA of genomic or synthetic origin, *i.e.*, a polymer of deoxyribonucleotide or ribonucleotide bases, respectively, read from the 5' (upstream) end to the 3' (downstream) end.

As used herein, the phrase "polynucleotide sequence" refers to the sequence of a polynucleotide molecule. The nomenclature for DNA bases as set forth at 37 CFR § 1.822 is used.

As used herein "heterologous" DNA is any polynucleotide sequence which is not naturally found next to the adjacent DNA. Heterologous DNA is often found in a DNA construct used for transformation. A p26 promoter operably linked to a reporter gene is an example of a heterologous DNA as the p26 promoter is naturally and normally associated with a p26 gene.

Promoters

As used herein, the term "promoter" refers to a polynucleotide molecule that in its native state is located upstream or 5' to a translational start codon of an open reading frame (or protein-coding region) and that is involved in recognition and binding of RNA polymerase II and other proteins (trans-acting transcription factors) to initiate transcription. A "plant promoter" is a native or non-native promoter that is functional in plant cells. Constitutive plant promoters are functional in most or all tissues of a plant throughout plant development. Any plant promoter can be used as a 5' regulatory element for modulating expression of a particular gene or genes operably associated thereto. When operably linked to a transcribable polynucleotide molecule, a promoter typically causes the transcribable polynucleotide molecule to be transcribed in a manner that is similar to that of which the promoter is normally associated. Plant promoters can include promoters produced through the manipulation of known promoters to produce artificial, chimeric, or hybrid promoters. Such promoters can also combine cis-elements from one or more promoters, for example, by adding a heterologous regulatory element to an active promoter with its own partial or complete regulatory elements. Thus, the design, construction, and use of chimeric or hybrid promoters comprising at least one cis-element of SEQ ID NOs: 1, 2, 3, or 4 for modulating the expression of operably linked polynucleotide sequences is encompassed by the present invention.

As used herein, the term "cis-element" refers to a cis-acting transcriptional regulatory element that confers an aspect of the overall control of gene expression. A cis-element may function to bind transcription factors, trans-acting protein factors that regulate transcription. Some cis-elements bind more than one transcription factor, and transcription factors may interact with different affinities with more than one cis-element. The promoters of the present invention desirably contain cis-elements that can confer or modulate gene expression. Cis-elements can be identified by a number of techniques, including deletion analysis, *i.e.*, deleting one or more nucleotides from the 5' end or internal to a promoter; DNA binding protein analysis using DNase I footprinting, methylation interference, electrophoresis mobility-shift assays, *in vivo* genomic footprinting by ligation-mediated PCR, and other conventional assays; or by DNA sequence similarity analysis with known cis-element motifs by conventional DNA sequence comparison methods. The fine structure of a cis-element can be further studied by mutagenesis (or substitution) of one or more nucleotides or by other conventional methods. Cis-elements can be obtained by chemical synthesis or by isolation from promoters that include such elements, and they can be synthesized with additional flanking nucleotides that contain useful restriction enzyme sites to facilitate subsequence manipulation.

In one embodiment, the promoters of the present invention comprise multiple cis-elements each of which confers a different aspect to the overall control of gene expression. In a preferred embodiment, cis-elements from the polynucleotide molecules of SEQ ID NOs: 1, 2, 3, and 4 are identified using computer programs designed specifically to identify cis-element, domains, or motifs within sequences. Cis-elements may either positively or negatively regulate gene expression, depending on the conditions. The present invention therefore encompasses cis-elements of the disclosed promoters.

As used herein, the phrase "substantially homologous" refers to polynucleotide molecules that generally demonstrate a substantial percent sequence identity with the promoters provided herein. Of particular interest are polynucleotide molecules wherein the polynucleotide molecules function in plants to direct transcription and have at least about 70% sequence identity, at least about 80% sequence identity, at least about 90% sequence identity, or even greater sequence identity, such as 98% or 99% sequence identity with the polynucleotide sequences of the promoters described herein. Polynucleotide molecules that are capable of regulating transcription of operably linked transcribable polynucleotide molecules and are substantially homologous to the polynucleotide sequences of the promoters provided herein are encompassed within the scope of this present invention.

As used herein, the phrase "percent sequence identity" refers to the percentage of identical nucleotides in a linear polynucleotide sequence of a reference polynucleotide molecule (or its complementary strand) as compared to a test polynucleotide molecule (or its complementary strand) when the two sequences are optimally aligned (with appropriate nucleotide insertions, deletions, or gaps totaling less than 20% of the reference sequence over the window of comparison). Optimal alignment of sequences for aligning a comparison window are well known to those skilled in the art and may be conducted by tools such as the local homology algorithm of Smith and Waterman, the homology alignment algorithm of Needleman and Wunsch, the search for similarity method of Pearson and Lipman, and preferably by computerized implementations of these algorithms such as GAP, BESTFIT, FASTA, and TFASTA available as part of the GCG® Wisconsin Package® (Accelrys Inc., San Diego, CA). An "identity fraction" for aligned segments of a test sequence and a reference sequence is the number of identical components which are shared by the two aligned sequences divided by the total number of components in the reference sequence segment, *i.e.*, the entire reference sequence or a smaller defined part of the reference sequence. Percent sequence identity is represented as the identity fraction times 100. The comparison of one or more polynucleotide sequences may be to a full-length polynucleotide sequence or a portion thereof, or to a longer polynucleotide sequence.

As used herein, the term "homology" refers to the level of similarity or percent identity between polynucleotide sequences in terms of percent nucleotide positional identity, *i.e.*, sequence similarity or identity. As used herein, the term homology also refers to the concept of similar functional properties among different polynucleotide molecules, e.g., promoters that have similar function may have homologous cis-elements. Polynucleotide molecules are homologous when under certain conditions they specifically hybridize to form a duplex molecule. Under these conditions, referred to as stringency conditions, one polynucleotide molecule can be used as a probe or primer to identify other polynucleotide molecules that share homology. The phrase "stringent conditions" is functionally defined with regard to the hybridization of a nucleic-acid probe to a target nucleic acid (*i.e.*, to a particular nucleic-acid sequence of interest) by the specific hybridization procedure discussed in *Molecular Cloning: A Laboratory Manual, 3rd edition Volumes 1, 2, and 3*. J.F. Sambrook, D.W. Russell, and N. Irwin, Cold Spring Harbor Laboratory Press, 2000 (referred to herein as Sambrook, *et al.*). Accordingly, the nucleotide sequences of the present invention may be used for their ability to selectively form duplex molecules with complementary stretches of polynucleotide molecule fragments. Depending on the application envisioned one would desire to employ varying

conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively high stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.15 M NaCl at temperatures of about 50°C to about 70°C. A high stringent condition, for example, is to wash the hybridization filter at least twice with high-stringency wash buffer (0.2 X SSC, 0.1% SDS, 65°C). Appropriate moderate stringency conditions that promote DNA hybridization, for example, 6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art. Additionally, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50°C to a high stringency of about 0.2 x SSC at 50°C. Additionally, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or either the temperature or the salt concentration may be held constant while the other variable is changed. Such selective conditions tolerate little mismatch between the probe and the template or target strand. Detection of polynucleotide molecules via hybridization is well known to those of skill in the art, and the teachings of U.S. Patents 4,965,188 and 5,176,995 are exemplary of the methods of hybridization analyses.

Homology can also be determined by computer programs that align polynucleotide sequences and estimate the ability of polynucleotide molecules to form duplex molecules under certain stringency conditions. Polynucleotide molecules from different sources that share a high degree of homology are referred to as "homologues".

Methods well known to one skilled in the art may be used to identify promoters of interest having activity similar to the promoters described herein. For example, cDNA libraries may be constructed using cells or tissues of interest and screened to identify genes having an expression pattern similar to that of the promoters described herein. The cDNA sequence for the identified gene may then be used to isolate the gene's promoter for further characterization. *See,* for example, U.S. Patents 6,096,950; 5,589,583; and 5,898,096; incorporated herein by reference. Alternately, transcriptional profiling or electronic northern techniques may be used to identify genes having an expression pattern similar to that of the promoters described herein. Once these genes have been identified, their promoters may be isolated for further characterization. *See*, for example, U.S. Patents 6,506,565 and 6,448,387, incorporated herein by reference. The electronic northern technique refers to a computer-based sequence analysis which allows sequences from multiple cDNA libraries to be

compared electronically based on parameters the researcher identifies including abundance in EST populations in multiple cDNA libraries, or exclusively to EST sets from one or combinations of libraries. The transcriptional profiling technique is a high-throughput method used for the systematic monitoring of gene expression profiles for thousands of genes. This DNA chip-based technology arrays thousands of cDNA sequences on a support surface. These arrays are simultaneously hybridized to a population of labeled cDNA probes prepared from RNA samples of different cell or tissue types, allowing direct comparative analysis of expression. This approach may be used for the isolation of regulatory sequences such as promoters associated with those genes.

In another embodiment, the promoter disclosed herein can be modified. Those skilled in the art can create promoters that have variations in the polynucleotide sequence. The polynucleotide sequences of the promoters of the present invention as shown in SEQ ID NOs: 1, 2, 3, or 4 may be modified or altered to enhance their control characteristics. One preferred method of alteration of a polynucleotide sequence is to use PCR to modify selected nucleotides or regions of sequences. These methods are well known to those of skill in the art. Sequences can be modified, for example by insertion, deletion, or replacement of template sequences in a PCR-based DNA modification approach. A "variant" is a promoter containing changes in which one or more nucleotides of an original promoter is deleted, added, and/or substituted, preferably while substantially maintaining promoter function. For example, one or more base pairs may be deleted from the 5' or 3' end of a promoter to produce a "truncated" promoter. One or more base pairs can also be inserted, deleted, or substituted internally to a promoter. In the case of a promoter fragment, variant promoters can include changes affecting the transcription of a minimal promoter to which it is operably linked. A minimal or basal promoter is a polynucleotide molecule that is capable of recruiting and binding the basal transcription machinery. One example of basal transcription machinery in eukaryotic cells is the RNA polymerase II complex and its accessory proteins. Variant promoters can be produced, for example, by standard DNA mutagenesis techniques or by chemically synthesizing the variant promoter or a portion thereof.

Novel chimeric promoters can be designed or engineered by a number of methods. Many promoters contain cis-elements that activate, enhance, or define the strength and/or specificity of the promoter. For example, promoters may contain "TATA" boxes defining the site of transcription initiation and other cis-elements located upstream of the transcription initiation site that modulate transcription levels. For example, a chimeric promoter may be produced by fusing a first promoter fragment containing the activator cis-element from one

promoter to a second promoter fragment containing the activator cis-element from another promoter; the resultant chimeric promoter may cause an increase in expression of an operably linked transcribable polynucleotide molecule. Promoters can be constructed such that promoter fragments or elements are operably linked, for example, by placing such a fragment upstream of a minimal promoter. The cis-elements and fragments of the present invention can be used for the construction of such chimeric promoters. Methods for construction of chimeric and variant promoters of the present invention include, but are not limited to, combining control elements of different promoters or duplicating portions or regions of a promoter (*see*, for example, U.S. Patents 4,990,607; 5,110,732; and 5,097,025, all of which are herein incorporated by reference). Those of skill in the art are familiar with the standard resource materials that describe specific conditions and procedures for the construction, manipulation, and isolation of macromolecules (*e.g.*, polynucleotide molecules, plasmids, etc.), as well as the generation of recombinant organisms and the screening and isolation of polynucleotide molecules.

In another embodiment, a promoter comprising the polynucleotide sequence shown in SEQ ID NO: 2 includes any length of said polynucleotide sequence that is capable of regulating an operably linked transcribable polynucleotide molecule. For example, the promoters as disclosed in SEQ ID NO: 2 may be truncated or portions deleted and still be capable of regulating transcription of an operably linked polynucleotide molecule. In a related embodiment, a cis-element of the disclosed promoters may confer a particular specificity such as conferring enhanced expression of operably linked polynucleotide molecules in certain tissues and therefore is also capable of regulating transcription of operably linked polynucleotide molecules. Consequently, any fragments, portions, or regions of the promoters comprising the polynucleotide sequence shown in SEQ ID NO: 2 can be used as regulatory polynucleotide molecules, including but not limited to cis-elements or motifs of the disclosed polynucleotide molecules. Substitutions, deletions, insertions, or any combination thereof can be combined to produce a final construct.

Polynucleotide Constructs

As used herein, the term "construct" refers to any recombinant polynucleotide molecule such as a plasmid, cosmid, virus, autonomously replicating polynucleotide molecule, phage, or linear or circular single-stranded or double-stranded DNA or RNA polynucleotide molecule, derived from any source, capable of genomic integration or autonomous replication, comprising a polynucleotide molecule where one or more polynucleotide molecule has been linked in a functionally operative manner.

As used herein, the phrase "operably linked" refers to a first polynucleotide molecule, such as a promoter, connected with a second transcribable polynucleotide molecule, such as a gene of interest, where the polynucleotide molecules are so arranged that the first polynucleotide molecule affects the function of the second polynucleotide molecule. Preferably, the two polynucleotide molecules are part of a single contiguous polynucleotide molecule and more preferably are adjacent. For example, a promoter is operably linked to a gene of interest if the promoter regulates or mediates transcription of the gene of interest in a cell.

As used herein, the phrase "transcribable polynucleotide molecule" refers to any polynucleotide molecule capable of being transcribed into a RNA molecule. Methods are known for introducing constructs into a cell in such a manner that the transcribable polynucleotide molecule is transcribed into a functional mRNA molecule that is translated and therefore expressed as a protein product. Constructs may also be constructed for transcription of antisense RNA molecules or other similar inhibitory RNA in order to inhibit expression of a specific RNA molecule of interest in a target host cell. For the practice of the present invention, conventional compositions and methods for preparing and using constructs and host cells are well known to one skilled in the art, *see*, for example, Sambrook, *et al.*

Constructs of the present invention would typically contain a promoter operably linked to a transcribable polynucleotide molecule operably linked to a 3' transcription termination polynucleotide molecule. In addition, constructs may include but are not limited to additional regulatory polynucleotide molecules from the 3'-untranslated region (3' UTR) of plant genes (*e.g.*, a 3' UTR to increase mRNA stability of the mRNA, such as the PI-II termination region of potato or the octopine or nopaline synthase 3' termination regions). Constructs may include but are not limited to the 5' untranslated regions (5' UTR) of an mRNA polynucleotide molecule which can play an important role in translation initiation and can also be a genetic component in a plant expression construct. For example, non-translated 5' leader polynucleotide molecules derived from heat shock protein genes have been demonstrated to enhance gene expression in plants (*see*, for example, U.S. Patents 5,659,122 and 5,362,865, herein incorporated by reference). These additional upstream and downstream regulatory polynucleotide molecules may be derived from a source that is native or heterologous with respect to the other elements present on the promoter construct.

Thus, constructs of the present invention comprise promoters such as those provided in SEQ ID NOs: 1, 2, 3, or 4 modified as described above, operatively linked to a transcribable polynucleotide molecule so as to direct transcription of said transcribable polynucleotide

molecule at a desired level or in a desired tissue or developmental pattern upon introduction of said construct into a plant cell. In some cases, the transcribable polynucleotide molecule comprises a protein-coding region of a gene, and the promoter provides for transcription of a functional mRNA molecule that is translated and expressed as a protein product. Constructs may also be constructed for transcription of antisense RNA molecules or other similar inhibitory RNA in order to inhibit expression of a specific RNA molecule of interest in a target host cell.

Exemplary transcribable polynucleotide molecules for incorporation into constructs of the present invention include, for example, DNA molecules or genes from a species other than the target gene species, or even genes that originate with or are present in the same species, but are incorporated into recipient cells by genetic engineering methods rather than classical reproduction or breeding techniques. Exogenous gene or genetic element is intended to refer to any gene or DNA molecule that is introduced into a recipient cell. The type of DNA included in the exogenous DNA can include DNA that is already present in the plant cell, DNA from another plant, DNA from a different organism, or a DNA generated externally, such as a DNA molecule containing an antisense message of a gene, or a DNA molecule encoding an artificial or modified version of a gene.

The promoters of the present invention can be incorporated into a construct using marker genes as described and tested in transient analyses that provide an indication of gene expression in stable plant systems. As used herein the phrase "marker gene" refers to any transcribable polynucleotide molecule whose expression can be screened for or scored in some way. Methods of testing for marker gene expression in transient assays are known to those of skill in the art. Transient expression of marker genes has been reported using a variety of plants, tissues, and DNA delivery systems. For example, types of transient analyses can include but are not limited to direct gene delivery via electroporation or particle bombardment of tissues in any transient plant assay using any plant species of interest. Such transient systems would include but are not limited to electroporation of protoplasts from a variety of tissue sources or particle bombardment of specific tissues of interest. The present invention encompasses the use of any transient expression system to evaluate promoters or promoter fragments operably linked to any transcribable polynucleotide molecules, including but not limited to selected reporter genes, marker genes, or genes of agronomic interest. Examples of plant tissues envisioned to test in transients via an appropriate delivery system would include but are not limited to leaf base tissues, callus, cotyledons, roots, endosperm, embryos, floral tissue, pollen, and epidermal tissue.

Any scorable or screenable marker gene can be used in a transient assay. Preferred marker genes for transient analyses of the promoters or promoter fragments of the present invention include a GUS gene (U.S. Patent 5,599,670, herein incorporated by reference) or a GFP gene (U.S. Patent 5,491,084, herein incorporated by reference). The constructs containing the promoters or promoter fragments operably linked to a marker gene are delivered to the tissues and the tissues are analyzed by the appropriate mechanism, depending on the marker. The quantitative or qualitative analyses are used as a tool to evaluate the potential expression profile of the promoters or promoter fragments when operatively linked to genes of agronomic interest in stable plants.

Thus, in one preferred embodiment, a polynucleotide molecule of the present invention as shown in SEQ ID NOs: 1, 2, 3, or 4 or fragments, variants, or derivatives thereof is incorporated into a construct such that a promoter of the present invention is operably linked to a transcribable polynucleotide molecule that provides for a selectable, screenable, or scorable marker. Markers for use in the practice of the present invention include, but are not limited to transcribable polynucleotide molecules encoding ß-glucuronidase (GUS), green fluorescent protein (GFP), luciferase (LUC), proteins that confer antibiotic resistance, or proteins that confer herbicide tolerance. Useful antibiotic resistance markers, including those encoding proteins conferring resistance to kanamycin (nptII), hygromycin B (aph IV), streptomycin or spectinomycin (aad, spec/strep), and gentamycin (aac3 and aacC4) are known in the art. Herbicides for which transgenic plant tolerance has been demonstrated and the method of the present invention can be applied, include but are not limited to: glyphosate, glufosinate, sulfonylureas, imidazolinones, bromoxynil, delapon, cyclohezanedione, protoporphyrionogen oxidase inhibitors, and isoxasflutole herbicides. Polynucleotide molecules encoding proteins involved in herbicide tolerance are known in the art, and include, but are not limited to a polynucleotide molecule encoding 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) described in U.S. Patents 5,627,061; 5,633,435; and 6,040,497; and *aroA* described in U.S. Patent 5,094,945 for glyphosate tolerance; a polynucleotide molecule encoding bromoxynil nitrilase (*Bxn*) described in U.S. Patent 4,810,648 for Bromoxynil tolerance; a polynucleotide molecule encoding phytoene desaturase (*crtI*) described in Misawa *et al.*, *Plant J.*, 4:833-840 (1993) and Misawa *et al.*, *Plant J.*, 6:481-489 (1994) for norflurazon tolerance; a polynucleotide molecule encoding acetohydroxyacid synthase (AHAS, *aka* ALS) described in Sathasiivan *et al.*, *Nucl. Acids Res.*, 18:2188-2193 (1990) for tolerance to sulfonylurea herbicides; and the *bar* gene

described in DeBlock, *et al.*, *EMBO J.*, 6:2513-2519 (1987) for glufosinate and bialaphos tolerance.

In one preferred embodiment, a polynucleotide molecule of the present invention as shown in SEQ ID NOs: 1, 2, 3, or 4 or fragments, variants, or derivatives thereof is incorporated into a construct such that a promoter of the present invention is operably linked to a transcribable polynucleotide molecule that is a gene of agronomic interest. As used herein, the phrase "gene of agronomic interest" refers to a transcribable polynucleotide molecule that includes but is not limited to a gene that provides a desirable characteristic associated with plant morphology, physiology, growth and development, yield, nutritional enhancement, disease or pest resistance, or environmental or chemical tolerance. The expression of a gene of agronomic interest is desirable in order to confer an agronomically important trait. A gene of agronomic interest that provides a beneficial agronomic trait to crop plants may be, for example, including, but not limited to genetic elements comprising herbicide resistance (U.S. Patents 5,633,435 and 5,463,175), increased yield (U.S. Patent 5,716,837), insect control (U.S. Patents 6,063,597; 6,063,756; 6,093,695; 5,942,664; and 6,110,464), fungal disease resistance (U.S. Patents 5,516,671; 5,773,696; 6,121,436; 6,316,407; and 6,506,962), virus resistance (U.S. Patents 5,304,730 and 6,013,864), nematode resistance (U.S. Patent 6,228,992), bacterial disease resistance (U.S. Patent 5,516,671), starch production (U.S. Patents 5,750,876 and 6,476,295), modified oils production (U.S. Patent 6,444,876), high oil production (U.S. Patents 5,608,149 and 6,476,295), modified fatty acid content (U.S. Patent 6,537,750), high protein production (U.S. Patent 6,380,466), fruit ripening (U.S. Patent 5,512,466), enhanced animal and human nutrition (U.S. Patents 5,985,605 and 6,171,640), biopolymers (U.S. Patent 5,958,745 and U.S. Patent Publication No. 2003/0028917), environmental stress resistance (U.S. Patent 6,072,103), pharmaceutical peptides (U.S. Patent 6,080,560), improved processing traits (U.S. Patent 6,476,295), improved digestibility (U.S. Patent 6,531,648), low raffinose (U.S. Patent 6,166,292), industrial enzyme production (U.S. Patent 5,543,576), improved flavor (U.S. Patent 6,011,199), nitrogen fixation (U.S. Patent 5,229,114), hybrid seed production (U.S. Patent 5,689,041), and biofuel production (U.S. Patent 5,998,700), the genetic elements and transgenes described in the patents listed above are herein incorporated by reference.

Useful nucleic acid sequences that can be combined with the promoter nucleic acid sequence of the present invention and provide improved end-product traits include, without limitation, those encoding seed storage proteins, fatty acid pathway enzymes, tocopherol biosynthetic enzymes, amino acid biosynthetic enzymes, and starch branching enzymes. A

discussion of exemplary heterologous DNAs useful for the modification of plant phenotypes may be found in, for example, U.S. Patents 6,194,636; 6,207,879; 6,232,526; 6,426,446; 6,429,357; 6,433,252; 6,437,217; 6,515,201; and 6,583,338 and PCT Publication WO 02/057471, each of which is specifically incorporated herein by reference in its entirety.

Preferred seed storage proteins include zeins (U.S. Patents 4,886,878; 4,885,357; 5,215,912; 5,589,616; 5,508,468; 5,939,599; 5,633,436; and 5,990,384; PCT Publications WO 90/01869, WO 91/13993, WO 92/14822, WO 93/08682, WO 94/20628, WO 97/28247, WO 98/26064, and WO 99/40209), 7S proteins (U.S. Patents 5,003,045 and 5,576,203), brazil nut protein (U.S. Patent 5,850,024), phenylalanine free proteins (PCT Publication WO 96/17064), albumin (PCT Publication WO 97/35023), b-conglycinin (PCT Publication WO 00/19839), 11S (U.S. Patent 6,107,051), alpha-hordothionin (U.S. Patents 5,885,802 and 5,88,5801), arcelin seed storage proteins (U.S. Patent 5,270,200), lectins (U.S. Patent 6,110,891), and glutenin (U.S. Patents 5,990,389 and 5,914,450) all of which are incorporated herein by reference.

Preferred fatty acid pathway enzymes include thioesterases (U.S. Patents 5,512,482; 5,530,186; 5,945,585; 5,639,790; 5,807,893; 5,955,650; 5,955,329; 5,759,829; 5,147,792; 5,304,481; 5,298,421; 5,344,771; and 5,760,206), diacylglycerol acyltransferases (U.S. Patent Publications 20030115632A1 and 20030028923A1), and desaturases (U.S. Patents 5,689,050; 5,663,068; 5,614,393; 5,856,157; 6,117,677; 6,043,411; 6,194,167; 5,705,391; 5,663,068; 5,552,306; 6,075,183; 6,051,754; 5,689,050; 5,789,220; 5,057,419; 5,654,402; 5,659,645; 6,100,091; 5,760,206; 6,172,106; 5,952,544; 5,866,789; 5,443,974; and 5,093,249) all of which are incorporated herein by reference.

Preferred tocopherol biosynthetic enzymes include *tyrA, slr1736, ATPT2, dxs, dxr, GGPPS, HPPD, GMT, MT1, tMT2, AANT1, slr 1737*, and an antisense construct for homogentisic acid dioxygenase (Kridl *et al., Seed Sci. Res.,* 1:209:219 (1991); Keegstra, *Cell,* 56(2):247-53 (1989); Nawrath *et al., Proc. Natl. Acad. Sci. USA,* 91:12760-12764 (1994); Xia *et al., J. Gen. Microbiol.*, 138:1309-1316 (1992); Lois *et al., Proc. Natl. Acad. Sci. USA,* 95 (5):2105-2110 (1998); Takahashi *et al., Proc. Natl. Acad. Sci. USA,* 95(17):9879-9884 (1998); Norris *et al., Plant Physiol.,* 117:1317-1323 (1998); Bartley and Scolnik, *Plant Physiol.,* 104:1469-1470 (1994); Smith *et al., Plant J.*, 11:83-92 (1997); WO 00/32757; WO 00/10380; Saint Guily *et al., Plant Physiol.*, 100(2):1069-1071 (1992); Sato *et al., J. DNA Res.,* 7(1):31-63 (2000)) all of which are incorporated herein by reference.

Preferred amino acid biosynthetic enzymes include anthranilate synthase (U.S. Patent 5,965,727 and PCT Publications WO 97/26366, WO 99/11800, and WO 99/49058),

tryptophan decarboxylase (PCT Publication WO 99/06581), threonine decarboxylase (U.S. Patents 5,534,421 and 5,942,660; PCT Publication WO 95/19442), threonine deaminase (PCT Publications WO 99/02656 and WO 98/55601), dihydrodipicolinic acid synthase (U.S. Patent 5,258,300), and aspartate kinase (U.S. Patents 5,367,110; 5,858,749; and 6,040,160) all of which are incorporated herein by reference.

Preferred starch branching enzymes include those set forth in U.S. Patents 6,232,122 and 6,147,279; and PCT Publication WO 97/22703, all of which are incorporated herein by reference.

Alternatively, a transcribable polynucleotide may be designed to down-regulate a specific nucleic acid sequence. This is typically accomplished by operably linking with a promoter, such as a promoter of the present invention, an exogenous DNA in an antisense orientation or a DNA designed such that a hairpin-forming RNA molecule is generated upon transcription. Gene suppression may be effective against a native plant gene associated with a trait, *e.g.*, to provide plants with reduced levels of a protein encoded by the native gene or with enhanced or reduced levels of an affected metabolite. For example, a promoter of the present invention may be operably linked to a heterologous DNA designed such that a hairpin-shaped RNA is formed for suppression of a native gene in dicotyledonous seed. The RNA could also be a catalytic RNA molecule (*i.e.*, a ribozyme) engineered to cleave a desired endogenous mRNA product. Thus, any polynucleotide molecule that encodes a protein or mRNA that expresses a phenotype or morphology change of interest is useful for the practice of the present invention.

As used herein "gene suppression" means any of the well-known methods for suppressing an RNA transcript or production of protein translated from an RNA transcript, including post-transcriptional gene suppression and transcriptional suppression. Post-transcriptional gene suppression is mediated by double-stranded RNA having homology to a gene targeted for suppression. Gene suppression by RNA transcribed from an exogenous DNA construct comprising an inverted repeat of at least part of a transcription unit is a common feature of gene suppression methods known as anti-sense suppression, co-suppression, and RNA interference. Transcriptional suppression can be mediated by a transcribed double-stranded RNA having homology to promoter DNA sequence to effect what is called promoter *trans*-suppression.

More particularly, post transcriptional gene suppression by inserting an exogenous DNA construct with anti-sense oriented DNA to regulate gene expression in plant cells is disclosed in U.S. Patents 5,107,065 and 5,759,829, each of which is incorporated herein by

reference in its entirety. Transgenic plants transformed using such anti-sense oriented DNA constructs for gene suppression can comprise DNA arranged as an inverted repeat, as disclosed by Redenbaugh *et al.*, in "Safety Assessment of Genetically Engineered Flavr Savr™ Tomato, CRC Press, Inc. (1992). Inverted repeat insertions can comprise a part or all of a T-DNA construct, *e.g.*, an inverted repeat of transcription terminator sequence.

Post transcriptional gene suppression by inserting an exogneous DNA construct with sense-oriented DNA to regulate gene expression in plants is disclosed in U.S. Patents 5,283,184 and 5,231,020, each of which is incorporated herein by reference.

Different types of exogenous DNA arrangements resulting in gene suppression are known to those of skill in the art and include but are not limited to the following. PCT Publication WO 94/01550 discloses DNA constructs where the anti-sense RNA was stabilized with a self-complementary 3' segment. Other double-stranded hairpin-forming elements in transcribed RNA are disclosed in PCT Publication No. 98/05770 where the anti-sense RNA is stabilized by hairpin forming repeats of poly(CG) nucleotides and U.S. Application Publication No. 2002/0048814A1 describes sense or anti-sense RNA stabilized by a poly(T)-poly(A) tail. U.S. Application Publication No. 2003/0018993A1 discloses sense or anti-sense RNA is stabilized by an inverted repeat of a subsequence of 3' untranslated region of the NOS gene. U.S. Application Publication No. 2003/0036197A1 describes an RNA stabilized by two complementary RNA regions having homology to a target sequence.

Gene silencing can also be effected by transcribing RNA from both a sense and an anti-sense oriented DNA, *e.g.*, as disclosed in U.S. Patent 5,107,065 and other examples as follows. U.S. Patent 6,326,193 discloses gene targeted DNA which is operably linked to opposing promoters. Sijen *et al.*, *The Plant Cell*, 8:2277-2294 (1996) discloses the use of constructs carrying inverted repeats of a cowpea mosaic virus gene in transgenic plants to mediate virus resistance. Such constructs for post transcriptional gene suppression in plants by double-stranded RNA are also disclosed in PCT Publication Nos. WO 99/53050, WO 99/49029, and U.S. Application Publication No. 2003/0175965A1, U.S. Application No.10/465,800, and U.S. Patent 6,506,559. *See*, also, U.S. Application No. 10/393,347 which discloses constructs and methods for simultaneously expressing one or more recombinant genes while simultaneously suppressing one or more native genes in a transgenic plant. *See*, also, U.S. Patent 6,448,473 which discloses multigene suppression vectors for use in plants. All of the above-described patents, applications and international publications disclosing materials and methods for post transcriptional gene suppression in plants are incorporated herein by reference.

Transcriptional suppression such as promoter *trans* suppression can be effected by expressing a DNA construct comprising a promoter operably linked to inverted repeats of promoter DNA for a target gene. Constructs useful for such gene suppression mediated by promoter *trans* suppression are disclosed by Mette *et al., The EMBO Journal*, 18(1):241-248, 1999 and by Mette *et al., The EMBO Journal*, 19(19):5194-5201, 2000), both of which are incorporated herein by reference.

The constructs of the present invention are generally double Ti plasmid border DNA constructs that have the right border (RB or AGRtu.RB) and left border (LB or AGRtu.LB) regions of the Ti plasmid isolated from *Agrobacterium tumefaciens* comprising a T-DNA, that along with transfer molecules provided by the *Agrobacterium* cells, permits the integration of the T-DNA into the genome of a plant cell. The constructs also contain the plasmid backbone DNA segments that provide replication function and antibiotic selection in bacterial cells, for example, an *E. coli* origin of replication such as *ori*322, a broad host range origin of replication such as *ori*V or *ori*Ri, and a coding region for a selectable marker such as Spec/Strp that encodes for Tn7 aminoglycoside adenyltransferase (*aadA*) conferring resistance to spectinomycin or streptomycin, or a gentamicin (Gm, Gent) selectable marker gene. For plant transformation, the host bacterial strain is often *Agrobacterium tumefaciens* ABI, C58, or LBA4404, however, other strains known to those skilled in the art of plant transformation can function in the present invention.

Transformed Plants And Plant Cells

As used herein, the term "transformed" refers to a cell, tissue, organ, or organism into which has been introduced a foreign polynucleotide molecule, such as a construct. The introduced polynucleotide molecule may be integrated into the genomic DNA of the recipient cell, tissue, organ, or organism such that the introduced polynucleotide molecule is inherited by subsequent progeny. A "transgenic" or "transformed" cell or organism also includes progeny of the cell or organism and progeny produced from a breeding program employing such a transgenic plant as a parent in a cross and exhibiting an altered phenotype resulting from the presence of a foreign polynucleotide molecule. A plant transformation construct containing a promoter of the present invention may be introduced into plants by any plant transformation method. Methods and materials for transforming plants by introducing a plant expression construct into a plant genome in the practice of this present invention can include any of the well-known and demonstrated methods including electroporation as illustrated in U.S. Patent 5,384,253; microprojectile bombardment as illustrated in U.S. Patents 5,015,580; 5,550,318; 5,538,880; 6,160,208; 6,399,861; and 6,403,865; *Agrobacterium*-mediated

transformation as illustrated in U.S. Patents 5,824,877; 5,591,616; 5,981,840; and 6,384,301; and protoplast transformation as illustrated in U.S. Patent 5,508,184, all of which are hereby incorporated by reference.

Methods for specifically transforming dicots are well known to those skilled in the art. Transformation and plant regeneration using these methods have been described for a number of crops including, but not limited to, cotton (*Gossypium hirsutum*), soybean (*Glycine max*), peanut (*Arachis hypogaea*), and members of the genus *Brassica*.

The transformed plants are analyzed for the presence of the genes of interest and the expression level and/or profile conferred by the promoters of the present invention. Those of skill in the art are aware of the numerous methods available for the analysis of transformed plants. For example, methods for plant analysis include, but are not limited to Southern blots or northern blots, PCR-based approaches, biochemical analyses, phenotypic screening methods, field evaluations, and immunodiagnostic assays.

The seeds of this present invention can be harvested from fertile transgenic plants and be used to grow progeny generations of transformed plants of this present invention including hybrid plant lines comprising the construct of this present invention and expressing a gene of agronomic interest.

The present invention also provides for parts of the plants of the present invention. Plant parts, without limitation, include seed, endosperm, ovule, and pollen. In a particularly preferred embodiment of the present invention, the plant part is a seed. It is understood that even after separation or isolation from other plant parts, the isolated or separated plant part may be contaminated with other plant parts. In a preferred aspect, the separated plant part is greater than 50% (w/w) of the separated material, more preferably, greater than 75% (w/w) of the separated material, and even more preferably greater than 90% (w/w) of the separated material. Plants or plant parts of the present invention generated by such methods may be processed into products using known techniques. Preferred products are meal, feedstock, and oil.

In another embodiment, the present invention provides a method of making a vegetable oil, comprising the steps of incorporating into the genome of an oilseed plant a promoter of the present invention operably linked to a transcribable polynucleotide molecule conferring altered oil content, growing the oilseed plant to produce oilseeds, and extracting the oil from the oilseed.

In another embodiment, the present invention provides a method of making a meal, comprising the steps of incorporating into the genome of an oilseed plant a promoter of the

present invention operably linked to a transcribable polynucleotide molecule conferring altered protein and/or micronutrient content, growing the oilseed plant to produce oilseeds, and producing the meal from the oilseed.

Methods to produce feed, meal, protein, and oil preparations are known in the art. *See*, for example, U.S. Patents 4,957,748; 5,100,679; 5,219,596; 5,936,069; 6,005,076; 6,146,669; and 6,156,227. In a preferred embodiment, the protein preparation is a high protein preparation. Such a high protein preparation preferably has a protein content of greater than about 5% w/v, more preferably greater than about 10% w/v, and even more preferably greater than about 15% w/v. In a preferred oil preparation, the oil preparation is a high oil preparation with an oil content derived from a plant or part thereof of the present invention of greater than about 5% w/v, more preferably greater than about 10% w/v, and even more preferably greater than about 15% w/v. In a preferred embodiment the oil preparation is a liquid and of a volume greater than 1, 5, 10, or 50 liters. The present invention provides for oil produced from plants of the present invention or generated by a method of the present invention. Such oil may be a minor or major component of any resultant product. Moreover, such oil may be blended with other oils. In a preferred embodiment, the oil produced from plants of the present invention or generated by a method of the present invention constitutes greater than about 0.5%, about 1%, about 5%, about 10%, about 25%, about 50%, about 75%, or about 90% by volume or weight of the oil component of any product. In another embodiment, the oil preparation may be blended and can constitute greater than about 10%, about 25%, about 35%, about 50%, or about 75% of the blend by volume. Oil produced from a plant of the present invention can be admixed with one or more organic solvents or petroleum distillates.

In a further embodiment, meal of the present invention may be blended with other meals. In a preferred embodiment, the meal produced from plants of the present invention or generated by a method of the present invention constitutes greater than about 0.5%, about 1%, about 5%, about 10%, about 25%, about 50%, about 75%, or about 90% by volume or weight of the meal component of any product. In another embodiment, the meal preparation may be blended and can constitute greater than about 10%, about 25%, about 35%, about 50%, or about 75% of the blend by volume.

The phrase "micronutrient content" means the amount of micronutrients, *i.e.*, vitamins A, E, K, tocopherols, tocotrienols, or carotenoids, within a seed expressed on a per weight basis.

The phrase "oil content" means oil level, which may be determined, for example, by low-resolution ^{1}H nuclear magnetic resonance (NMR) (Tiwari *et al.*, *JAOCS*, 51:104-109, 1974 or Rubel, *JAOCS*, 71:1057-1062, 1994) or near infrared transmittance (NIT) spectroscopy (Orman *et al.*, *JAOCS*, 69(10):1036-1038, 1992 and Patrick *et al.*, *JAOCS*, 74(3):273-276, 1997).

The phrase "protein quality" means the level of one or more essential amino acids, whether free or incorporated in protein, namely histidine, isoleucine, leucine, lysine, methionine, cysteine, phenylalanine, tyrosine, threonine, tryptophan, and valine.

The following examples are included to demonstrate preferred embodiments of the present invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples that follow represent techniques discovered by the inventors to function well in the practice of the present invention. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments that are disclosed and still obtain a like or similar result without departing from the spirit and scope of the present invention, therefore all matter set forth or shown in the accompanying drawings is to be interpreted as illustrative and not in a limiting sense.

Example 1

This example sets forth the isolation and characterization of the nucleic acid sequences for the plant promoters p26, p63, p63tr and pBAN.

p26

A cDNA clone, designated clone 26, was identified from *Arabidopsis thaliana* using a cDNA-AFLP procedure. Briefly, SMART cDNA libraries were prepared from mRNA isolated from *Arabidopsis* according to manufacturer's instructions (Clontech Laboratories, Palo Alto, CA). The mRNA was isolated from open flowers (inflorescence), stem, whole seedling, and developing seed harvested at 4, 7, 10, 13, or 18 days after flowering (DAF). Five hundred micrograms of amplified SMART cDNA was used for AFLP analysis using the Gibco-BRL small genome AFLP II Kit (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. The resulting bands were visualized after electrophoresis through 6% acrylamide/8M Urea sequencing gels. A single band, designated band 26, visible only in the lane derived from cDNA from 4 DAF developing seed tissue, was extracted from the gel. The DNA was eluted by placing the acrylamide band in 50 µl of TE (10mM Tris-HCl

(pH 8.0); 1 mM EDTA) and allowing the band to elute at ambient temperature overnight. After a brief centrifugation, 0.75 µl of supernatant was used as a template source for a PCR amplification. Twenty microliters of Preamp primer mix 2 (Gibco-BRL small genome AFLP II Kit, Invitrogen), 2.5 µl 10X PCR buffer with 15 mM MgCl (PE Applied Biosystems Foster City, CA), 2.5 units AmpliTaq DNA Polymerase (PE Applied Biosystems), and 2 µl water were added to the DNA template. Amplification conditions were as follows: 25 cycles of 94°C for 30 seconds, 56°C for 60 seconds, and 72°C for 2 minutes followed by 1 cycle of 72°C for 2 minutes. The resulting DNA fragment was subcloned into pCR2.1Topo (Invitrogen) according to manufacturers' instructions. This cDNA clone 26 was hybridized to virtual Northerns that were prepared from SMART cDNA libraries, according to the manufacturer's instructions, using the SMART PCR cDNA Synthesis Kit (BD Biosciences, Clontech, Palo Alto, CA). Clone 26 was characterized as being expressed in early stages of seed development. The sequence of clone 26 was determined using standard sequencing methodologies as set forth by PE Applied Biosystems BigDye terminator v.3.0 (PE Applied Biosystems, Foster City, CA).

The entire sequence of clone 26 was then used as a query for a BLAST search against public and proprietary genomic DNA databases. A single EST clone, LIB3176-P1-K1-C12, was identified from the search. The clone had homology to a palmitoyl-protein thioesterase in *Arabidopsis thaliana* (GenBank protein_id BAA97167; gi:8809616). The entire genomic sequence of clone 26 was contained within the P1 clone MQL5 (SEQ ID NO: 5, GenBank Accession AB018117; gi:3702735).

To prepare a genomic library from *Arabidopsis*, genomic DNA was isolated using a modification of a genomic DNA isolation protocol (Dellaporta *et al.*, *Plant Molecular Biology Reporter,* 1:19-21, 1983). Soil or plate grown *Arabidopsis* seedlings were harvested and kept frozen in liquid nitrogen until extraction. The tissue was ground to a fine powder using a mortar and pestle while keeping the tissue frozen with liquid nitrogen. The ground tissue was transferred to a Waring blender containing 200 ml of cold (0°C) DNA extraction buffer (350 mM sorbitol; 100 mM Tris; 5 mM EDTA; pH to 7.5 with HCl; sodium bisulfite (3.8 mg/ml) added just before use, and homogenized at high speed for 30-60 seconds. The homogenate was filtered through a layer of cheesecloth and collected in a centrifuge bottle. The samples were centrifuged at 2500xg for 20 minutes. The supernatant and any loose green material was discarded. The pellet was then resuspended in 1.25 ml DNA extraction buffer and transferred to a 50 ml polypropylene tube. Then 1.75 ml nuclei lysis buffer (200 mM Tris; 50 mM EDTA; 2 M NaCl; 2% CTAB (Hexadecyltrimethyl-Ammonium Bromide,

Sigma, St. Louis, MO); pH to 7.5 with HCl), and 0.6 ml of 5% (w/v) sarkosyl was added. The tubes were mixed gently, and the samples were incubated at 65°C for 20 minutes. An equal volume of chloroform:isoamyl alcohol (24:1) was added and the tubes were mixed gently. The tubes were then centrifuged at 2500xg for 15 minutes, and the resulting supernatant was transferred to a clean tube. An equal volume of ice-cold isopropanol was poured onto the sample, and the sample was inverted several times until a precipitate formed. The precipitate was removed from the solution using a glass pipette and residual alcohol removed by allowing the precipitate to air dry for 2-5 minutes. The precipitate was resuspended in 400 µl TE buffer (10mM Tris-HCl pH 8.0, 1 mM EDTA).

Arabidopsis genomic DNA as prepared above was used to prepare Genome Walker libraries (Clontech) according to the manufacturer's instructions. The p26 promoter sequence (SEQ ID NO: 2) was isolated from the libraries according to the manufacturer's instructions. Primers were designed based upon the sequence of the P1 clone MQL5 (SEQ ID NO: 5). The following were used as the primers for the PCR reaction:

Clone 26 GSP1

5'-ATCGGCAACTCCATTTCCAATTTCTC-3' (SEQ ID NO: 7)

and Clone 26 GSP2

5'-TAGCATCCCTAGCATTAGAACATTGAG-3' (SEQ ID NO: 8).

After an initial PCR using the Genome Walker libraries as template and primers GSP1 and AP1 (supplied by the manufacturer), a second amplification was performed using the first round amplification product as template and primers GSP2 and AP2 (supplied by the manufacturer).

The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid was named pMON69803 and contained the sequence of the promoter p26 (SEQ ID NO: 2). The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems, (Foster City, CA).

pBAN

The promoter sequence for pBAN was identified from a BLAST search of the coding region of the BANYULS gene (GenBank AF092912, Devic *et.al.*, *The Plant Journal,* 19(4):387-398, 1999) against an *Arabidopsis* database in GenBank, the NIH genetic sequence database containing an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research*, 30(1):17-20, 2002). The search identified BAC clone T13M11 (GenBank

AC005882, [SEQ ID NO: 6]) that contains the BANYULS coding region in antisense orientation. The following primers were designed to amplify the sequences corresponding to T12M11 base pairs 44629-45570 from *Arabidopsis* genomic DNA prepared as described above.

BAN+1500

5'-GTTTGATAACTCGTCTCTTG-3' (SEQ ID NO: 9)

and BAN GSP1

5'-GGTGTGTGTAAGAGTCTGGTCC-3' (SEQ ID NO: 10)

The reaction conditions for the PCR followed a protocol recommended by the enzyme manufacturer (PE Applied Biosystems, Foster City, CA). The Banyuls promoter was isolated from *Arabidopsis* genomic DNA using 30 nanomoles each of primers Ban+1500 and Ban GSP1, 10 micromoles each of dATP, dCTP, dGTP, and TTP, 2.5 units of AmpliTaq Gold in 1X Opti-Prime™ Buffer 3 (Stratagene, La Jolla, CA). After an initial incubation at 95°C for 10 minutes, 30 cycles of PCR were performed with 92°C for 30 seconds, 56°C for 30 seconds, and 72°C for 2 minutes followed by 1 cycle of 72°C for 7 minutes.

The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid was named pBAN1.

To add an NcoI site at the predicted ATG start codon of the Banyuls gene an additional PCR reaction was performed. pBAN1 was used as template DNA with the following primers:

Ban-Nco

5'-CCATGGTTGTACTTTTGAAATTACAGAG-3' (SEQ ID NO: 11)

and Ban+1500

5'-GTTTGATAACTCGTCTCTTG-3' (SEQ ID NO: 9)

The reaction conditions for the second PCR reaction followed a protocol recommended by the enzyme manufacturer (PE Applied Biosystems, Foster City, CA). Approximately 10 nanograms of pBAN1 is amplified using 30 nanomoles each of primers Ban+1500 and Ban-Nco, 10 micromoles each of dATP, dCTP, dGTP, and TTP, 2.5 units of AmpliTaq Gold in 1X Opti-Prime™ Buffer 3 (Stratagene). After an initial incubation at 95°C for 10 minutes, 30 cycles of PCR were performed with 92°C for 30 seconds, 56°C for 30 seconds, and 72°C for 2 minutes followed by 1 cycle of 72°C for 7 minutes.

The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's

instructions. The resulting plasmid was named pMON69809 and contained the sequence of the promoter pBAN (SEQ ID NO: 1). The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems.

p63

A cDNA clone, designated clone 63, was identified from *Arabidopsis thaliana* using cDNA-AFLP. SMART cDNA libraries were prepared from mRNA isolated from *Arabidopsis* open flowers (inflorescence), stem, whole seedling, and developing seed harvested at 4, 7, 10, 13, or 18 days after flowering (DAF), according to manufacturer's instructions (Clontech). Five hundred micrograms of amplified SMART cDNA was used for AFLP analysis using the Gibco-BRL small genome AFLP II kit and following manufacturer's instructions (Invitrogen). The resulting bands were visualized after electrophoresis through 6% acrylamide/8M Urea sequencing gels. A single band, designated band 63, visible only in the lane derived from cDNA from 4 DAF developing seed tissue, was extracted from the gel. The DNA was eluted by placing the acrylamide band in 50 µl of TE (10mM Tris-HCl (pH 8.0); 1 mM EDTA) and allowing the band to elute at ambient temperature overnight. After a brief centrifugation, 0.75 µl of supernatant was used as the template source for a PCR amplification. Twenty microliters of Preamp primer mix 2 (from the Gibco-BRL small genome AFLP II Kit), 2.5 µl 10X PCR buffer with 15 mM $MgCl_2$ (PE Applied Biosystems), 2.5 units AmpliTaq DNA Polymerase (PE Applied Biosystems), and 2 µl water were added to the DNA template. Amplification conditions were as follows: 25 cycles of 94°C for 30 seconds, 56°C for 60 seconds, and 72°C for 2 minutes followed by 1 cycle of 72°C for 2 minutes. The resulting DNA fragment was subcloned into pCR2.1 Topo (Invitrogen) according to manufacturers' instructions. This cDNA clone 63 was hybridized to virtual Northerns that were prepared from SMART cDNA libraries, according to the manufacturer's instructions, using the SMART PCR cDNA Synthesis Kit (BD Biosciences, Clontech, Palo Alto, CA). Clone 63 was characterized as being expressed in early stages of seed development. The sequence of clone 63 was determined using standard sequencing methodologies as set forth by PE Applied Biosystems BigDye terminator v.3.0 (PE Applied Biosystems, Foster City, CA).

The entire sequence of the clone 63 was used as a query for a BLAST search against public and proprietary databases. Public database searches indicated that clone 63 was annotated as a putative protein. The entire genomic sequence of clone 63 was contained

within the BAC clone T25N20 (Choi *et al.*, *Weeds World,* 2:17-20, 1995), which was then obtained from the Arabidopsis Biological Resource Center (Columbus, OH).

An overnight culture of *E. coli* containing the BAC clone T25N20 was grown from a single colony in LB broth (10% bacto-tryptone, 5% yeast extract, and 10% NaCl with kanamycin (25 mg/L) and ampicillin (100 mg/L)), containing 12.5 mg/liter chloramphenicol at 37°C with vigorous shaking until late exponential or early stationary phase. Clone 63 (p63) was then isolated from the BAC clone T25N20. The cells were collected via centrifugation resulting pellet was resuspended in 0.2 ml buffer (50mM glucose; 10mM EDTA; 25mM Tris pH 8.0; 5 mg/ml lysozyme) and incubated on ice for 5 minutes, followed by the addition of 0.4 ml of 0.2 N NaOH; 1% SDS solution. The tube was mixed gently and incubated on ice for 5 minutes, followed by the addition of 0.3 ml of 3 M potassium acetate. The tube was mixed gently and then frozen at minus 80°C for 15 minutes. The debris was pelleted by centrifugation at 20,000xg for 15 minutes, and 0.75 ml of the resulting supernatant was transferred to a new tube. Isopropanol (0.45 ml) was added and the mixture was incubated at minus 80°C for 15 minutes. DNA was pelleted by centrifugation at 20,000xg for 5 minutes. The pellet was rinsed with 1 ml of cold 70% ethanol, then dried on the bench for at least 15 minutes prior to being resuspended in 40 µl TE buffer.

The following primers were used to PCR amplify p63 from BAC T25N20:

p63-Nco

5'-CCATGGTTATTCAAGTGACCACAG-3' (SEQ ID NO: 12)

and p63-fwd3

5'-CGTGTTGAGGTGAGAGG-3' (SEQ ID NO: 13)

The conditions for the PCR reaction followed a protocol recommended by the enzyme manufacturer (PE Applied Biosystems, Foster City, CA). The p63 sequence was amplified using 1.5 µl of T25N20 as template, 30 nanomoles each of the primers p63-Nco and p63-fwd3, 10 micromoles each of dATP, dCTP, dGTP, and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene). After an initial incubation at 95°C for 10 minutes, 28 cycles of PCR were performed with 94°C for 15 seconds, 62°C for 10 seconds, 52°C for 10 seconds, and 72°C for 3 minutes followed by 1 cycle of 72°C for 7 minutes.

The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid was named pMON69811 and contained the sequence of

the promoter p63 (SEQ ID NO: 3). The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems.

Example 2

This example describes the construction of the vectors used for *Arabidopsis* transformation.

pMON69802

An 1861 base pair (bp) fragment containing the *E. coli uid*A gene (GUS) was removed from the donor plasmid pCGN10906 by digestion with EcoRI. The fragment was isolated from an agarose gel using the QiaGel Purification kit (Qiagen) according to the manufacturer's instructions. The purified DNA was eluted from the column using 30 µl of Buffer EB (10 mM Tris-Cl pH 8.5). New restriction endonuclease sites were added to the *E. coli uid*A gene using primers:

Gus 5'
5'-AGGCGGCGCCTAAACCATGGTCCGTCCTGTAGAAACCCC-3' (SEQ ID NO: 14)
and Gus 3'
5'-AGTCGACTCATTGTTTGCCTCCCTGCTGCGGTTTTTCAC-3' (SEQ ID NO: 15).

The purified fragment (0.5 µl) was used as the template for the following PCR amplification. Thirty nanomoles each of primers Gus 5' (SEQ ID NO: 14) and GUS 3' (SEQ ID NO: 15), 10 micromoles each of dATP, dCTP, dGTP, and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene) were added to the DNA template. Amplification conditions were as follows: 25 cycles of 94°C for 30 seconds, 56°C for 60 seconds, and 72°C for 2 minutes followed by 1 cycle of 72°C for 2 minutes. The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid was named pMON65400. The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems.

A 1278 bp fragment containing the napin 3' UTR was removed from the vector pCGN 7770 by digestion with SalI and Asp718I. An 1861 base pair fragment, containing the *E. coli uid*A gene, was removed from the vector pMON65400 by sequential digestion with BstXI and SalI. Prior to SalI digestion and gel purification, the BstXI overhang was blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene). Both fragments were ligated into the vector pCGN8541, which had been digested with Asp718I and SwaI. The resulting plasmid, containing the *E. coli uid*A gene and the napin 3' UTR, was named

pMON69802 (Figure 1). The nucleic acid sequence was determined using known methodology and confirmed the integrity of the cloning junctions.

pMON69804 (p26::GUS)

A 1007 bp fragment containing the p26 sequence was removed from pMON69803 by digestion with HindIII and NcoI. The fragment was ligated into pMON69802, which had also been digested with HindIII and NcoI. The resulting plasmid, containing the p26 promoter driving the *E. coli uid*A gene and with the napin 3' UTR was named pMON69804 (Figure 2). The nucleic acid sequence was determined using known methodology and confirmed the integrity of the cloning junctions. This vector was used in the subsequent transformation of *Arabidopsis*.

pMON69815 (pBAN::GUS)

A 947bp BstXI-NcoI fragment containing the pBAN promoter sequence was removed from pMON69809 by sequential digestion with BstXI followed by and NcoI. Prior to NcoI digestion and gel purification, the BstXI overhang was blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene). The fragment was ligated into pMON69802, which had been sequentially digested with HindIII and NcoI. Prior to NcoI digestion and gel purification, the HindIII overhang from pMON69802 was blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene). The resulting plasmid, containing the pBAN promoter driving the *E. coli uid*A gene and with the napin 3' UTR, was named pMON69815 (Figure 3). The nucleic acid sequence was determined using known methodology and confirmed the integrity of the cloning junctions. This vector was used in the subsequent transformation of *Arabidopsis*.

pMON69812 (p63::GUS)

A 2031 base pair fragment containing the p63 sequence was cut out of pMON69811 by sequential digestion with BstXI and NcoI. Prior to NcoI digestion and gel purification, the BsatXI overhang was blunt ended using Pfu polymerase according to manufacturer's instructions (Stratagene). The fragment was ligated into pMON69802, which had been digested with HindIII, and also had the resulting overhang blunt ended with Pfu polymerase. The resulting plasmid, containing the p63 promoter driving the *E. coli uid*A gene and with the napin 3' UTR was named pMON69812 (Figure 4). The nucleic acid sequence was determined using known methodology and confirmed the integrity of the cloning junctions. This vector was used in the subsequent transformation of *Arabidopsis*.

Example 3

This example describes the transformation and subsequent regeneration of transgenic *Arabidopsis* plants expressing a heterologous gene of interest.

Arabidopsis plants were grown by sowing seeds onto 4 inch pots containing reverse osmosis water (ROW) saturated MetroMix 200 (The Scotts Company, Columbus, OH). The plants were vernalized by placing the pots in a covered flat, in a growth chamber at 4-7°C, 8 hours light/day for 4-7 days. The flats were transferred to a growth chamber at 22°C, 55% relative humidity, and 16 hours light/day at an average intensity of 160-200 μEinstein/s/m^2. The cover was lifted and slid back 1 inch after germination, and then was removed when the true leaves had formed. The plants were bottom watered, as needed, with ROW until 2-3 weeks after germination. Plants were then bottom watered, as needed, with Plantex 15-15-18 solution (Plantex Corporation Ottawa, Canada) at 50 ppm N_2. Pots were thinned so that 1 plant remained per pot at 2-3 weeks after germination. Once the plants began to bolt, the primary inflorescence was trimmed to encourage the growth of axillary bolts.

Transgenic *Arabidopsis thaliana* plants were obtained as described by Bent *et al.*, *Science,* 265:1856-1860, 1994 or Bechtold *et al.*, *C.R.Acad.Sci, Life Sciences*, 316:1194-1199, 1993. Cultures of *Agrobacterium tumefaciens* strain ABI containing one of the transformation vectors pMON69804, pMON69812, or pMON69815 were grown overnight in LB (10% bacto-tryptone, 5% yeast extract, and 10% NaCl with kanamycin (75 mg/L), chloramphenicol (25 mg/L), and spectinomycin (100 mg/L)). The bacterial culture was centrifuged and resuspended in 5% sucrose + .05% Silwet-77 solution. The aerial portions of whole *Arabidopsis thaliana* plants (at about 5-7 weeks of age) were immersed in the resulting solution for 2-3 seconds. The excess solution was removed by blotting the plants on paper towels. The dipped plants were placed on their side in a covered flat and transferred to a growth chamber at 19°C. After 16 to 24 hours the dome was removed and the plants were set upright. When plants had reached maturity, water was withheld for 2-7 days prior to seed harvest. Harvested seed was passed through a stainless steel mesh screen (40 holes/inch) to remove debris. The harvested seed was stored in paper coin envelopes at room temperature until analysis.

The harvested seeds described above were sown onto flats containing ROW saturated MetroMix 200 (The Scotts Company). The plants were vernalized and germinated as described above. After true leaves had emerged, the aerial portion of the seedlings were sprayed with a solution containing a 1:200 dilution of Finale herbicide (The Scotts Company). Approximately 1 week after the first application, the plants were sprayed a second time. Up

to 16 Finale resistant seedlings were transplanted to 2¼ inch pots, one seedling per pot, containing MetroMix 200 and were grown under the conditions described above until the initial siliques that had formed began to desiccate. Tissue (rosette leaf, cauline leaf, stem, flowers, floral buds, and developing siliques) was removed from each T1 plant for subsequent histochemical staining.

Example 4

Expression of β-glucuronidase was analyzed in *Arabidopsis thaliana* plants transformed with pMON69815, pMON69812, or pMON69804 using histochemical staining. Tissues, prepared as described in Example 3, were incubated for approximately 24 hours at 37°C in a solution containing 50 mM $NaPO_4$ (pH 7.2); 100 μM potassium ferricyanide; 100 μM potassium ferrocyanide, 0.03% Triton X-100; 20% methanol and 2.5 mg/ml 5-bromo-4-chloro-3-indoyl glucuronic acid (X-gluc). In some cases the potassium ferricyanide, potassium ferrocyanide, and methanol were omitted from the staining solution. The stained tissue was cleared of chlorophyll by an overnight incubation in 70% ethanol/30% H_2O at 37°C. Stained tissues were photographed immediately or transferred to a solution of 70% ethanol/30% glycerol (v/v) and stored at 4°C until photographed. The results, summarized in Table 1 below, show that 4 out of the 11 individual T1 plants tested from pMON69804, (p26::GUS), showed GUS expression in the seed. For pMON69812, (p63::GUS), 13 of the 15 lines tested had expression in the seed. For pMON69815, (pBAN::GUS), 12 out of 12 lines tested had expression in the seed.

Table 1

Construct	Promoter	# of Lines Tested	Lines with Seed Expression
pMON69804	p26	11	4
pMON69815	pBan	12	12
pMON69812	p63	15	13

To examine the developmental stage at which the promoters were active, seeds from the independent lines that were positive for GUS expression in the T1 generation (described above) were sown onto pots containing ROW saturated MetroMix 200. The plants were vernalized, in a growth chamber, at 4-7°C and 8 hours of light/day for 4-7 days. The plants were transferred to a growth chamber at 22°C, 55% relative humidity, and 16 hours of light/day at an average intensity of 160-200 μEinstein/s/m^2. The plants were bottom watered, as needed, with ROW until well established, generally 2-3 weeks after germination. Plants

were then bottom watered, as needed, with Plantex 15-15-18 at 50ppm N_2. Pots were thinned so that 1 plant remained per 2¼ inch pot at 2-3 weeks after germination. At least 10 plants from each line were stained, as described above, at each time point. Visual observations of the GUS expression patterns were recorded. Qualitative expression was compared to the positive control plants containing a pNapin::GUS construct (labeled as 10908) and to the null segregants which served as the negative control plants. The results are shown in Table 2.

Expression driven by the napin promoter is detected from 7-18 days after flowering (daf). Expression driven by the p26 promoter, pBAN promoter, and the p63 promoter, was detected from 5-10, 1-14, and 4-14 daf, respectively. Hence, expression of all three promoters was detected earlier than that of the napin promoter.

Table 2

		Days after flowering												
Line	**promoter**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**10**	**13**	**14**	**18**
10908-7	napin	-	-	ND	ND	ND	-	ND	+	ND	+	ND	+	+
10908-10	napin	-	-	-	ND	ND	-	ND	+	ND	+	ND	+	ND
10908-16	napin	-	-	-	-	ND	-	-	-	ND	+	ND	-	-
69804-13	p26	-	-	-	-	-	ND	ND	+	ND	-	ND	-	-
69804-14	p26	-	-	-	-	-	+	ND	+	+	+	ND	-	-
69804-7	p26	-	-	-	ND	ND	+	ND	+	ND	-	ND	-	-
69815-2	pBan	-	+	+	+	ND	+	ND	+	ND	+	+	ND	-
69815-9	pBan	-	+	+	+	ND	+	ND	+	ND	+	ND	ND	-
69815-14	pBan	-	+	+	+	ND	+	ND	+	ND	+	ND	+	-
69812-4	p63	-	-	-	-	-	+	ND	+	ND	+	ND	+	-
69812-9	p63	-	-	-	-	ND	-	ND	-	ND	+	ND	+	-
69812-16	p63	-	-	-	-	ND	-	ND	-	ND	+	ND	+	-
69812-13	p63	ND	ND	ND	ND	+	+	ND	ND	ND	ND	ND	ND	ND

ND: Not Determined

Example 5

This example describes the vector construction and transformation of soy plants with reporter genes driven by promoters of the present invention.

Vector Construction

pMON82350 (p63::GUS)

A 3859-base pair fragment containing p63 (SEQ ID NO: 3), the *Escherichia coli* uidA gene and the napin 3' UTR was removed from the vector pMON65415 by digestion with NotI and Sse8387I. The fragment was ligated in between the 3' UTR from the pea rbsc E9 gene

and octopine T-DNA left border sequence in the vector pMON65448, which had been digested with NotI and Sse8387I. The vector pMON65448 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence, in between which are the promoter, 5' UTR and first intron from the Arabidopsis act7 gene driving the expression of a CP4 EPSP synthase gene containing a CTP, linked to a synthetic EPSP synthase coding region and the 3' UTR from the pea rbsc E9 gene. The resulting plasmid was named pMON82350 (Figure 5). The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

pMON82350 was transformed into soy via *Agrobacterium*-mediated transformation as described by Martinell *et al.*, U.S. Patent 6,384,301. Transformed plant tissues are collected and stained as described in Example 4. Gus expression is detected in the seed.

Example 6

This example describes the transformation and regeneration of canola plants with the heterologous gene of interest.

Vector Construction

a. pMON65422

To analyze the expression of p63 in Canola, a binary vector was constructed. A 3796 base pair fragment containing 703 base pairs of the p63 promoter (p63tr) (SEQ ID NO: 4), the *E. coli uid*A gene and the napin 3' untranslated region (UTR) was removed from pMON69812 by digestion with HindIII and Asp718I. Prior to gel purification the HindIII and the Asp718I overhangs were blunt ended using Pfu polymerase (Stratagene). The fragment was ligated into a PmeI digested pMON70650 backbone. The vector pMON70650 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence, with a 35S promoter from the Figwort Mosaic Virus (FMV) between the two T-DNA borders, driving the expression of a chimeric EPSP synthase gene containing a chloroplast targeting sequence from the *Arabidopsis* EPSP synthase gene (GenBank identifier number gi:16272) linked to a synthetic EPSP synthase coding region (U.S. Patent 5,633,435) and the 3' untranslated region from the pea rbcS E9 gene. Additionally, pMON70650 contains recognition sites for cre recombinase. The recombinase sites are 5' of the FMV promoter and 3' of the E9 3'. The resulting plasmid was designated pMON65422. DNA sequence analysis confirmed the integrity of the cloning junctions.

b. pMON65428

To analyze expression of p26 in Canola, a binary vector was derived from the vector pCGN11123. The vector pCGN11123 contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence, with an FMV-35S promoter, between the two T-DNA borders, driving the expression of a chimeric EPSP synthase gene containing a chloroplast targeting sequence from the *Arabidopsis* EPSP synthase gene (gi:16272) linked to a synthetic EPSP synthase coding region (U.S. Patent 5,633,435), the 3' UTR from the pea rbcS E9 gene, and recognition sites for cre recombinase.

A 745 base pair fragment of DNA was amplified from pMON70650 using the following primers:

CP4-Dra

5'-ACTTCACTTGAGCGGAAGCCATAG-3' (SEQ ID NO: 16)

and CP4-Kpn

5'-TTTAAAACAATGGCGCAAGTTAGCAG-3' (SEQ ID NO: 17).

The CP4-Dra primer causes a single nucleotide substitution in the 5' UTR of EPSP synthase that eliminates an NcoI restriction site. The NcoI site was removed to facilitate later cloning. The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo according to manufacturer's instructions (Invitrogen). The resulting plasmid was named pDMRUEZ033297. The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems.

A 737 base pair fragment containing the altered portion of the EPSP synthase gene was removed from pDMRUEZ033297 by digestion with KpnI and Dra I and ligated in place of the same size fragment of the vector pCGN11123. The resulting plasmid was named pDMRUEZ033298. A 3149 base pair fragment containing the *E. coli uid*A gene, and the napin 3' UTR was removed from pMON69802 by digestion with NotI and Asp718I. Prior to gel purification, the NotI and Asp718I overhangs were blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene). The fragment was ligated into a NotI digested pDMRUEZ033298 vector backbone. Prior to ligation, the NotI overhangs were blunt ended using Pfu polymerase (Stratagene) according to the manufacturer's instructions. The resulting plasmid contains the nopaline T-DNA right border sequence, the *E. coli uid*A gene, and the napin 3' UTR followed by an expression cassette with an FMV-35S promoter, between the two T-DNA borders, driving the expression of a chimeric EPSP synthase gene containing a chloroplast targeting sequence from the *Arabidopsis* EPSP synthase gene

(GenBank identifier number gi:16272) linked to a synthetic EPSP synthase coding region (U.S. Patent 5,633,435), the 3' UTR from the pea rbcS E9 gene, and recognition sites for cre recombinase, followed by the octopine T-DNA left border sequence. This plasmid was named pMON65424. DNA sequence analysis confirmed the integrity of the cloning junctions.

An 820 base pair fragment containing the p26 promoter sequence was removed from pMON69804 by digestion with SmaI and NcoI (Figure 2, [SEQ ID NO: 2]). The fragment was ligated into a PmeI-NcoI digested pMON65424 vector backbone. The resulting plasmid was named pMON65428 (Figure 7). The nucleic acid sequence was determined using known methodology and confirmed the integrity of the cloning junctions.

Canola Transformation

The vectors pMON65428 and pMON65422 are introduced into *Agrobacterium tumefaciens* strain ABI for transformation into *Brassica napus*. Canola plants are transformed using the protocol described by Moloney and Radke in U.S. Patent 5,720,871. Briefly, seeds of *Brassica napus* cv Ebony are planted in 2 inch pots containing Metro Mix 350 (The Scotts Company, Columbus, OH). The plants are grown in a growth chamber at 24°C, and a 16/8 hour photoperiod, with light intensity of 400 μEm^{-2} sec^{-1} (HID lamps). After 2-1/2 weeks, the plants are transplanted into 6 inch pots and grown in a growth chamber at 15/10°C day/night temperature, 16/8 hour photoperiod, light intensity of 800 μEm^{-2} sec^{-1} (HID lamps).

Four terminal internodes from plants just prior to bolting or in the process of bolting but before flowering are removed and surface sterilized in 70% v/v ethanol for 1 minute, 2% w/v sodium hypochlorite for 20 minutes and rinsing 3 times with sterile deionized water. Six to seven stem segments are cut into 5 mm discs, maintaining orientation of basal end.

The *Agrobacterium* culture used to transform Canola is grown overnight on a rotator shaker at 24°C in 2 mls of Luria Broth, LB, (10% bacto-tryptone, 5% yeast extract, and 10% NaCl) containing 50 mg/l kanamycin, 24 mg/l chloramphenicol, and 100 mg/l spectinomycin. A 1:10 dilution is made in MS media (Murashige and Skoog, *Physiol.Plant.*, 15:473-497, 1962) giving approximately $9x10^8$ cells per ml. The stem discs (explants) are inoculated with 1.0 ml of *Agrobacterium* and the excess is aspirated from the explants.

The explants are placed basal side down in petri plates containing media comprising 1/10 MS salts, B5 vitamins (1% inositol; 0.1% thiamine HCl; 0.01% nicotinic acid; 0.01% pyridoxine-HCl), 3% sucrose, 0.8% agar, pH 5.7, 1.0 mg/l 6-benzyladenine (BA). The plates

are layered with 1.5 ml of media containing MS salts, B5 vitamins, 3% sucrose, pH 5.7, 4.0 mg/l p-chlorophenoxyacetic acid, 0.005 mg/l kinetin and covered with sterile filter paper.

Following a 2 to 3 day co-culture, the explants are transferred to deep dish petri plates containing MS salts, B5 vitamins, 3% sucrose, 0.8% agar, pH 5.7, 1 mg/l BA, 500 mg/l carbenicillin, 50mg/1 cefotaxime, 200 mg/l kanamycin, or 175 mg/l gentamycin for selection. Seven explants are placed on each plate. After 3 weeks they are transferred to fresh media, 5 explants per plate. The explants are cultured in a growth room at 25°C, continuous light (Cool White).

The transformed plants are grown in a growth chamber at 22°C in a 16-8 hours light-dark cycle with light intensity of 220 $\mu Em^{-2}s^{-1}$ for several weeks before transferring to the greenhouse. Plants are maintained in a greenhouse under standard conditions. Developing seed is harvested at various stages after pollination and stored at minus 70°C. Mature seed is collected and stored under controlled conditions consisting of about 17°C and 30% humidity.

Up to 5 siliques are harvested from individual R0 plants at several time points after pollination. Siliques are scored with an 18 gauge needle to allow the staining solution to contact the developing seed. The siliques are incubated for approximately 24 hours at 37°C in a solution containing 50 mM $NaPO_4$ (pH 7.2); 100 μM potassium ferricyanide; 100 μM potassium ferrocyanide, 0.03% Triton X-100; 20% methanol and 2.5 mg/ml 5-bromo-4-chloro-3-indoyl glucuronic acid (X-gluc). The stained tissue is cleared of chlorophyll by an overnight incubation in 70% ethanol/30% H_2O at 37°C. Stained tissues are photographed immediately or transferred to a solution of 70% ethanol/30% glycerol (v/v) and stored at 4°C until photographed. Samples were scored positive (+) or negative (-) for blue color.

Six out of 10 lines transformed with pMON65428 have detectable levels of activity in seeds from at least one time point. Ten out of 10 lines transformed with pMON65422 have detectable levels of activity in seeds from at least one time point. No staining is observed in seeds harvested from nontransgenic control plants. The data is illustrated in the tables below.

p26 Expression in Developing Canola Seed

	Days After Pollination									
Construct	*3*	*6*	*9*	*12*	*15*	*20*	*25*	*30*	*35*	*40*
pMON65422	+	+	+	+	+	+	+	+	+	+
pMON65422	-	-	-	+	-	-	+	+	+	+
pMON65422	+	+	+	+	+	+	+	+	-	-
pMON65422	+		+	+	+	+	+	-	+	+
pMON65422	-	-	-	-	-	+	+	+	+	+
pMON65422	+	+	+	+	+	+	+	+	+	+
pMON65422	+	+	+	+	+	+	+	+	+	+
pMON65422	-	-	-	-	-	+	+	-	+	-
pMON65422	-	+	+	+	+	+	+	+	+	+
pMON65422	-	+	+	+	+	+	+	+	+	+
Control	-	-	-	-	-	-	-	-	-	-

p63 Expression in Developing Canola Seed

	Days After Pollination									
Construct	*3*	*6*	*9*	*12*	*15*	*20*	*25*	*30*	*35*	*40*
pMON65422	+	+	+	+	+	+	+	+	+	+
pMON65422	-	-	-	+	-	-	+	+	+	+
pMON65422	+	+	+	+	+	+	+	+	-	-
pMON65422	+		+	+	+	+	+	-	+	+
pMON65422	-	-	-	-	-	+	+	+	+	+
pMON65422	+	+	+	+	+	+	+	+	+	+
pMON65422	+	+	+	+	+	+	+	+	+	+
pMON65422	-	-	-	-	-	+	+	-	+	-
pMON65422	-	+	+	+	+	+	+	+	+	+
pMON65422	-	+	+	+	+	+	+	+	+	+
Control	-	-	-	-	-	-	-	-	-	-

Having illustrated and described the principles of the present invention, it should be apparent to persons skilled in the art that the present invention can be modified in arrangement and detail without departing from such principles. We claim all modifications that are within the spirit and scope of the appended claims. All publications and published patent documents cited in this specification are incorporated herein by reference to the same extent as if each individual publication or patent application is specifically and individually indicated to be incorporated by reference.